US005276741A

United States Patent [19]

Aragon

[11] Patent Number: 5,276,741

[45] Date of Patent: Jan. 4, 1994

[54] FUZZY STRING MATCHER

[75] Inventor: David B. Aragon, Berkeley, Calif.

[73] Assignee: TRW Financial Systems & Services, Inc., Cleveland, Ohio

[21] Appl. No.: 701,383

[22] Filed: May 16, 1991

[51] Int. Cl.^{5} G06F 9/72
[52] U.S. Cl. 382/40; 382/14; 395/61; 395/63
[58] Field of Search 382/40, 14, 15, 57; 364/518; 395/61, 63; G06K 9/72

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,058,795	11/1977	Balm	382/40
4,328,561	5/1982	Convis et al.	382/40
4,418,423	11/1983	Tsuji et al.	382/40
4,654,875	3/1987	Srihari et al.	382/40
4,979,227	12/1990	Mittelbach et al.	382/14
5,062,143	10/1991	Schmitt	382/40

FOREIGN PATENT DOCUMENTS

0031493	7/1981	European Pat. Off. .
1602591	11/1981	United Kingdom .

OTHER PUBLICATIONS

Tucker/Deaton, libsell, TRW, 1989, manual pages.
"A Comparison of Some Novel and Traditional Lexical Distance Metrics for Spelling Correction"; K. Kukich; pp. 309-313; Jul. 1990.
"A Very Fast Associative Method for the Recognition and Correction of Misspelt Words, Based on Redundant Hash Addressing"; T. Kohonen et al.; pp. 807-809; Nov. 1978.
European Search Report, 2 pp.; mailed Apr. 15, 1993.
Levenshtein, Binary Codes Capable of Correction In-

sertions, Deletions and Reverals, Sov. Phys. Dockl., vol. 10, Feb. 1966.
Kohonen, Self–Organizational and Assocative Memory, Springer–Verlag, 2nd ed., 1988.
Young/Fu, Handbook of Pattern Recognition and Image Processing, Academic Press, 1986.
Haines/Hecht–Nielsen, A BAM w/Increased Information Storage Capacity, ICNN Pro., Jul. 1988, vol. I, pp. 181–190.
Kukich, Variations on Back–Propagation Name Recognition Net, Proc. USPS Adv. Tech. Conf., 1988, pp. 722–735.
Cherassky/Vassilas, Performance of Back Propagation Networks for Assoc. Database Retrieval, Proc. IJCNN–89, vol. II, Jun. 1989.
Lampinen/Oja, Fast Self–Organization by Probing Algorithm, Proc. of IJCNN–89, Jun. 1989, vol. 4.
Bradburn, review of Lampinen/Oja in Neural Network Review, vol. 4, 1990.
Kohonen, Self–Organizing Map, Proceedings of IEEE, vol. 78, No. 9, Sep. 1990.
Venta/Kohonen, Content–Addressable Software Method for Emulation of Neural Networks, Proc. of ICNN, Jul. 1988.

Primary Examiner—Michael T. Razavi
Assistant Examiner—Yon J. Conso
Attorney, Agent, or Firm—Sheldon & Mak

[57] **ABSTRACT**

An apparatus and method useful in correcting character strings misread by a character recognition engine. The misread string is compared to character strings contained in a lexicon of valid character strings and the similarity between the strings measured. The valid strings can be ranked by degree of similarity should further selection be required to determine the appropriate value of the misread character string.

48 Claims, 8 Drawing Sheets